Claims

1. A method of docking a ligand to a protein comprising:

performing a pre-docking conformational search to generate multiple solution conformations of the ligand;

generating a binding site image of the protein, said binding site image comprising multiple hot spots;

matching hot spots of the binding site image to atoms in at least one solution conformation of the multiple solution conformations of the ligand to obtain at least one ligand position relative to the protein in a ligand-protein complex formation; and

optimizing the at least one ligand position while allowing translation, orientation and rotatable bonds of the ligand to vary, and while holding the protein fixed.

2. The method of claim 1, wherein said performing the pre-docking conformational search comprises creating a database of the multiple solution conformations and storing said three-dimensional database for subsequent use by said matching.

3. The method of claim 2, wherein said database of the multiple solution conformations comprises a conformational database of a combinatorial library.

4. The method of claim 1, wherein said performing the pre-docking conformational search comprises:

randomly generating a plurality of uniformly distributed conformations of the ligand;

minimizing a strain of each conformation of the plurality of uniformly distributed conformations;

using the strain and a solvent accessible surface area of each conformation to rank the conformations; and

clustering the conformations and retaining a desired number of top clusters of conformations, said retained number of top clusters of conformations comprising said multiple solution conformations of the ligand.

5. The method of claim 1, wherein said generating the binding site image includes at least one of creating a list of apolar hot spots identifying points in the binding site that are favorable for an apolar atom to bind, and generating a list of polar hot spots identifying points in the binding site that are favorable for a hydrogen bond donor or acceptor to bind.

6. The method of claim 5, wherein said generating the binding site image further comprises:

placing a grid around the binding site of the protein;

determining a hot spot search volume using said grid;

determining hot spots using a grid-like search of the hot spot search volume; and

for each type of hot spot, clustering the hot spots and retaining a desired number of clusters of hot spots with best scores, said desired number of clusters comprising said multiple hot spots to be employed by said matching.

7. The method of claim 1, wherein said matching comprises:

SUB B3

matching atoms of the at least one solution conformation to appropriate hot spots of the protein by positioning the at least one solution conformation as a rigid body into the binding site image;

defining a match, said match determining a unique rigid body transformation; and

using the unique rigid body transformation to place the at least one solution conformation of the ligand into the binding site of the protein.

SUB A16

8. The method of claim 7, wherein said determining the unique rigid body transformation comprises determining the unique rigid body transformation that minimizes:

$$I(R, T) = \sum_{j=1}^{3} \left| H_j - RA_j - T \right|^2$$

where:

H_j = a j^{th} hot spot of the protein;

A_j = a j^{th} atom of the at least one solution conformation;

R = a 3×3 rotation matrix; and

T = a translation vector.

9. The method of claim 1, wherein said optimizing comprises optimizing multiple protein-ligand complex formations, said optimizing comprising:

eliminating each ligand position having a predetermined percentage of ligand atoms with a steric clash;

ranking remaining ligand positions using an atom pairwise score with a desired atom score cutoff;

after ranking, clustering the ligand positions and selecting a top number n of ligand positions; and

optimizing each ligand position of the n positions, allowing the translation, rotation and rotatable bonds of the ligand to vary.

10. The method of claim 9, wherein said optimizing comprises optimizing each ligand position of the n positions using a BFGS optimization algorithm with a simple atom pairwise score, allowing the translation, rotation and rotatable bonds of the ligand to vary.

11. A system for docking a ligand to a protein comprising:

means for performing a pre-docking conformational search to generate multiple solution conformations of the ligand;

means for generating a binding site image of the protein, said binding site image comprising multiple hot spots;

means for matching hot spots of the binding site image to atoms in at least one solution conformation of the multiple solution conformations of the ligand to obtain at least one ligand position relative to the protein; and

means for optimizing the at least one ligand position while allowing translation, orientation and rotatable bonds of the ligand to vary, and while holding the protein fixed.

Sub C1

12. The system of claim 11, wherein said means for performing the pre-docking conformational search comprises means for creating a database of the multiple solution conformations and for storing said three-dimensional database for subsequent use by said matching.

13. The system of claim 12, wherein said database of the multiple solution conformations comprises a conformational database of a combinatorial library.

14. The system of claim 11, wherein said means for performing the pre-docking conformational search comprises:

means for randomly generating a plurality of uniformly distributed conformations of the ligand;

means for minimizing a strain of each conformation of the plurality of uniformly distributed conformations;

means for using the strain and a solvent accessible surface area of each conformation to rank the conformations; and

means for clustering the conformations and retaining a desired number of top clusters of conformations, said retained number of top clusters of conformations comprising said multiple solution conformations of the ligand.

15. The system of claim 11, wherein said means for generating the binding site image includes at least one of means for creating a list of apolar hot spots identifying points in the binding site that are favorable for an apolar atom to bind, and means for generating a list of polar hot spots identifying points in the binding site that are favorable for a hydrogen bond donor or acceptor to bind.

16. The system of claim 15, wherein said means for generating the binding site image further comprises:

means for placing a grid around the binding site of the protein;

means for determining a hot spot search volume using said grid;

means for determining hot spots using a grid-like search of the hot spot search volume; and

for each type of hot spot, means for clustering the hot spots and for retaining a desired number of clusters of hot spots with best scores, said desired number of clusters comprising said multiple hot spots to be employed by said matching.

17. The system of claim 11, wherein said means for matching comprises:

means for matching atoms of the at least one solution conformation to appropriate hot spots of the protein by positioning the at least one solution conformation as a rigid body into the binding site image;

means for defining a match, said match determining a unique rigid body transformation; and

means for using the unique rigid body transformation to place the at least one solution conformation of the ligand into the binding site of the protein.

18. The system of claim 17, wherein said determining the unique rigid body transformation comprises determining the unique rigid body transformation that minimizes:

$$I(R, T) = \sum_{j=1}^{3} \left| H_j - RA_j - T \right|^2$$

where:

H_j = a j^{th} hot spot of the protein;

A_j = a j^{th} atom of the at least one solution conformation;

R = a 3×3 rotation matrix; and

T = a translation vector.

19. The system of claim 11, wherein said means for optimizing comprises means for optimizing multiple protein-ligand complex formations, said means for optimizing comprising:

means for eliminating each ligand position having a predetermined percentage of ligand atoms with a steric clash;

means for ranking remaining ligand positions using an atom pairwise score with a desired atom score cutoff;

after ranking, means for clustering the ligand positions and selecting a top number n of ligand positions; and

means for optimizing each ligand position of the n positions, allowing the translation, rotation and rotatable bonds of the ligand to vary.

20. The system of claim 19, wherein said means for optimizing comprises means for optimizing each ligand position of the n positions using a BFGS optimization algorithm with a simple atom pairwise score, allowing the translation, rotation and rotatable bonds of the ligand to vary.

21. At least one program storage device readable by a machine, tangibly embodying at least one program of instructions executable by the machine to perform a method of docking a ligand to a protein, comprising:

performing a pre-docking conformational search to generate multiple solution conformations of the ligand;

generating a binding site image of the protein, said binding site image comprising multiple hot spots;

matching hot spots of the binding site image to atoms in at least one solution conformation of the multiple solution conformations of the ligand to obtain at least one ligand position relative to the protein; and

optimizing the at least one ligand position while allowing translation, orientation and rotatable bonds of the ligand to vary, and while holding the protein fixed.

Sub C1B

22. The at least one program storage device of claim 21, wherein said performing the pre-docking conformational search comprises creating a database of the multiple solution conformations and storing said three-dimensional database for subsequent use by said matching.

23. The at least one program storage device of claim 22, wherein said database of the multiple solution conformations comprises a conformational database of a combinatorial library.

24. The at least one program storage device of claim 21, wherein said performing the pre-docking conformational search comprises:

randomly generating a plurality of uniformly distributed conformations of the ligand;

minimizing a strain and a solvent accessible surface area of each conformation of the plurality of uniformly distributed conformations;

using the strain of each conformation to rank the conformations; and

clustering the conformations and retaining a desired number of top clusters of conformations, said retained number of top clusters of conformations comprising said multiple solution conformations of the ligand.

25. The at least one program storage device of claim 21, wherein said generating the binding site image includes at least one of creating a list of apolar hot spots identifying points in the binding site that are favorable for an apolar atom to bind, and generating a list of polar hot spots identifying points in the binding site that are favorable for a hydrogen bond donor or acceptor to bind.

26. The at least one program storage device of claim 25, wherein said generating the binding site image further comprises:

placing a grid around the binding site of the protein;

determining a hot spot search volume using said grid;

determining hot spots using a grid-like search of the hot spot search volume; and

for each type of hot spot, clustering the hot spots and retaining a desired number of clusters of hot spots with best scores, said desired number of clusters comprising said multiple hot spots to be employed by said matching.

27. The at least one program storage device of claim 21, wherein said matching comprises:

matching atoms of the at least one solution conformation to appropriate hot spots of the protein by positioning the at least one solution conformation as a rigid body into the binding site image;

defining a match, said match determining a unique rigid body transformation; and

using the unique rigid body transformation to place the at least one solution conformation of the ligand into the binding site of the protein.

28. The at least one program storage device of claim 27, wherein said determining the unique rigid body

transformation comprises determining the unique rigid body transformation that minimizes:

$$I(R, T) = \sum_{j=1}^{3} \left| H_j - RA_j - T \right|^2$$

where:

H_j = a j^{th} hot spot of the protein;

A_j = a j^{th} atom of the at least one solution conformation;

R = a 3×3 rotation matrix; and

T = a translation vector.

29. The at least one program storage device of claim 21, wherein said optimizing comprises optimizing multiple protein-ligand complex formations, said optimizing comprising:

eliminating each ligand position having a predetermined percentage of ligand atoms with a steric clash;

ranking remaining ligand positions using an atom pairwise score with a desired atom score cutoff;

after ranking, clustering the ligand positions and selecting a top number n of ligand positions; and

optimizing each ligand position of the n positions, allowing the translation, rotation and rotatable bonds of the ligand to vary.

30. The at least one program storage device of claim 29, wherein said optimizing comprises optimizing each ligand position of the n positions using a BFGS optimization algorithm with a simple atom pairwise score, allowing the translation, rotation and rotatable bonds of the ligand to vary.

* * * * *